3D Surface Reconstruction and Analysis in Automated Apple Stem-End/Calyx Identification

L. Jiang, B. Zhu, X. Cheng, Y. Luo, Y. Tao

Abstract. *Machine vision methods are widely used in apple defect detection and quality grading applications. Currently, 2D near-infrared (NIR) imaging technology is used to detect apple defects based on the difference in image intensity of defects from normal apple tissue. However, it is difficult to accurately differentiate an apple's stem-end/calyx from a true defect due to their similar 2D NIR images, which presents a major technical challenge to the successful application of this machine vision technology. In this research, we used a novel two-step 3D data analysis strategy to differentiate apple stem-ends/calyxes from true defects according to their different 3D shape information. In the first step, a 2D NIR imaging was extended to a 3D reconstruction using a shape-from-shading (SFS) approach. After successfully obtaining 3D information, a quadratic facet model was introduced to conduct the 3D concave shape fitting such that the identification of apple stem-ends and calyxes could be achieved based on their different 3D structures. Significant improvement in terms of the detection rate could be obtained based on 3D shape fitting in comparison to the traditional 2D intensity fitting approach. Samples of the reconstructed 3D apple surface maps as well as the identified stem-ends/calyxes were shown in the results, and an overall 90.15% detection rate was achieved, compared to the 58.62% detection rate of the traditional 2D intensity fitting approach.*

Keywords. *3D reconstruction, Apples, Automated detection, Calyx, Facet model, Near-infrared, Shape-from-shading, Stem end.*

Machine vision technology plays an important role in the apple industry by transforming the traditional apple-by-apple visual inspection to automated on-line sorting and grading. Researchers have investigated the properties of different light spectra from short to long wavelengths, and employed those light spectra into machine vision-based apple grading and sorting systems. Some of the systems (Good Fruit Growers, 1993) have been successfully used in the industry. Shahin et al. (2002) used an X-ray line scanner to acquire X-ray images of apples. Then spatial and transform image features were extracted and fed into separate artificial neural network (ANN) classifiers in order to distinguish between different bruise types on the apple surface. Yang and Marchant (1996) employed a charge-coupled device (CCD) monochromatic video camera to capture apple images in a lighting chamber, followed by a flooding algorithm to coarsely segment out apple defects such as bruises, insect bites, and scabs. Subsequently, an active contour model was applied to refine the segmentation in order to improve the localization and size accuracy of the detected blemishes. Leemans et al. (1999) chose a three-color CCD camera to acquire color images of bi-color apples. A method to segment defects, based on a Bayesian classification process, was then used. Among all the studied spectra, Brown et al. (1974) showed that in the near-infrared (NIR) image range between 700 and 2000 nm, there was less reflectance in bruised areas than in unbruised areas on apples. Since then, as an effective, while low-cost, imaging technology, NIR-based approaches for apple inspection have been intensively studied (Wen and Tao, 1998a; Tao and Wen, 1999; Wen and Tao, 1999; Li et al., 2002; Zhu et al., 2007a; Zhu et al., 2007c).

In most machine vision-based automated apple grading and sorting systems, it is important to identify apple stem-ends and calyxes in apple images because these images often exhibit patterns and intensity values that are similar to defects and result in false alarms during defect sorting. In addition, stem-end/calyx identification is necessary for estimating the apple firmness because the location of the stem-end and calyx must be known if an efficient firmness measurement device is to be perfected (Throop et al., 2001). To solve this problem, Wen and Tao (2000) built a dual-camera imaging system that incorporated an NIR camera (700 to 1000 nm) and a middle-infrared (MIR) camera (3.4 to 5 μm) to identify apple stem-ends and calyxes. They discovered that unlike a traditional NIR camera that is sensitive to stem-ends, calyxes, and defects, the MIR camera was only sensitive to apple stem-ends and calyxes. Based on this fact, the true defects were easily extracted by comparing the NIR and MIR images. Although a very high detection rate could be obtained based on the aforementioned NIR/MIR system, the cost of an MIR imaging device was too high to be accepted by the industry. Unay and Gosselin (2004) developed a two-cascaded-classifier approach to localize stem-ends and calyxes of Jonagold apples. First, an ANN was used to extract candidate

Submitted for review in February 2008 as manuscript number IET 7386; approved for publication by the Information & Electrical Technologies Division of ASABE in August 2009.

The authors are **Lu Jiang, ASABE Member Engineer,** Doctoral Student, **Bin Zhu,** Former Doctoral Student, **Xuemei Cheng,** Former Doctoral Student, **Yaguang Luo,** Scientist, Environmental Microbial and Food Safety Lab, USDA ARS, Beltsville, Maryland; and **Yang Tao,** Professor, Bio-imaging and Machine Vision Laboratory, Fischell Department of Bioengineering, University of Maryland, College Park, Maryland. **Corresponding author:** Yang Tao, University of Maryland, 1427 Agriculture and Animal Science Bldg. College Park, MD 20742; phone: 301-405-1189; fax: 301-314-9023; e-mail: ytao@umd.edu.

  ISSN 0001-2351

objects. A nearest-neighbor classifier was then applied to discriminate stem-ends and calyxes from other candidates. Penman (2001) utilized blue linear light sources and a standard color video camera to detect apple reflection patterns, which were formed by light stripes. Because the reflection patterns were shape and orientation dependent, it was possible to identify the location of stem-end and calyx regions. Throop et al. (2001) tested two conveyers for automatic apple orientation. Bi-rollers were used in the conveyer system to prevent the stem-end/calyx from showing in the camera's field of view in a mechanical way. Bennedsen et al. (2005) set up an experimental machine vision system to locate apple surface defects while eliminating other non-defect dark areas. The basic idea of their method was to rotate apples in front of the camera so that multiple images were acquired. Dark areas in these images, which kept the same position relative to the apple during the rotation, were considered defects, while other dark areas, whose positions kept changing, were classified as non-defects, such as stem-ends and calyxes.

Because of the similarity in the images of apple stem-ends/calyxes and defects, it is generally difficult to distinguish them based on their 2D information such as shape and image intensity. However, apple stem-ends and calyxes have special 3D characteristics, including bowl-shaped concaves. Considerable effort has been made in the area of apple 3D property analysis. Zion et al. (1995) developed a fast computerized method to detect bruises based on magnetic resonance imaging (MRI) images of apples. This approach has the potential to be expanded to 3D imaging and image analysis if the computation time of the 3D reconstruction algorithm can be reduced such that on-line processing requirements are met. Yang (1996) used structured lighting to detect stem-ends and calyxes. A set of evenly spaced parallel light stripes were projected onto the apple surfaces simultaneously. Generally, the stripes on convex apple surfaces had a parallel and parabolic pattern. However, when the stem-ends/calyxes came into view, this pattern was disturbed, and sharp change/broken stripes were observed around concave areas. Based on such disturbances, localization of stem-ends/calyxes could be achieved. The limitation of such an approach was that when the stem-end or calyx was oriented in the same direction as the stripe light source, the deformation of stripes was not obtainable. To compensate for the gradient reflectance of the curved surface of an apple, Wen and Tao (1998b) introduced a brightness-invariant image segmentation method for on-line fruit defect detection; Tao (1996) developed a spherical transform algorithm that converted a 3D image to a 2D image by means of compensating the intensity gradient on curved objects, such as apples, so that the intensity distribution became almost uniform after the transform. A preservation transform was also introduced in order to extract defects with the intensity below background level. Jing and Tao (1999) designed a laser range imaging system for real-time high-resolution 3-D shape reconstruction of images of poultry meat. A single laser line was projected onto samples, and two cameras were synchronized by a conveyer belt and used to continuously grab the laser profile images. Based on the triangular relationship among the laser line, camera view angle, and sample thickness, the 3D shapes of samples were precisely rebuilt. Since this was not a sample-dependent approach, it had great potential to be applied to apple 3D shape recovery.

The objective of this research was to develop an apple 3D shape recovery/analysis-based scheme for the reconstruction of 3D apple surfaces and the efficient identification of stem-ends and calyxes. In this study, a novel two-step 3D data analysis strategy was introduced to identify apple stem-ends and calyxes. In the first step, the Lambertian model was employed to evaluate the reflectance map of the apple surface, and then a shape-from-shading method was applied to rebuild the 3D apple surface based on a simplified human perception model (Zhu et al., 2005; Pentland, 1989). After the 3D reconstruction, a quadratic facet model was used to explain, and hence detect, the 3D concave shape of the apple stem-end/calyx.

Methods

Image Acquisition and Materials

The machine vision system (fig. 1) for apple inspection consisted of a computer-controlled image acquisition module and an NIR sensing system, which was a Hitachi KP-MI CCD monochromatic camera with a C-mount lens and 16 mm focal length and a Corrion 700 nm interference long-pass filter. The wavelength range of this system was from 700 nm to 1000 nm. Image resolution was 1.09 mm pixel^{-1}, the shutter speed was 1/250 s, and the online imaging system acquired the images at a rate of 30 frames s^{-1}.

A lighting chamber made by Agri-Tech, Inc., was used to provide uniform illumination for the infrared sensor. The 120 (W) × 100 (L) × 25 (H) cm chamber was made of lattice-patterned sheet metal, and the V-shaped interior surface of the chamber was painted flat white to provide diffuse light reflection and eliminate shadows. Lighting was provided by ten warm-white fluorescent lamps (GE SPX30 fluorescent linear lamp with 32 W and 110 VAC power supply) arranged uniformly around a V-shaped surface right above the conveyor and used to provide uniform illumination for the infrared sensor (Wen and Tao, 2000; Cheng et al., 2003). A conveyor comprised of rollers at a speed of 10 cm s^{-1} was used to keep the apples separate and to rotate each apple freely in order to make every apple side available to the NIR camera during multiple exposures.

A total of 203 NIR Golden Delicious apple images were acquired, of which 63 were samples without stem-ends/calyxes facing the camera, while 140 had stem-ends/calyxes showing in the image. The sample apples were placed on the conveyer manually at random orientations.

Preprocessing

An original NIR image of apples is shown in figure 2a. The intensity of the background varies and is relatively darker than the apples. In order to perform the 3D reconstruction, the non-uniform background has to be removed. In addition, each apple is a region of interest (ROI) and needs to be extracted individually. In this study, based on the total of 203 NIR images of Golden Delicious apples, a single threshold T_1 was used to coarsely segment apples from the background. Because some dark areas within apples could also be removed during the thresholding, a two-step morphological operation was then employed to refine the segmentation (Zhu et al., 2005). First, a hole-fill operation was performed, which filled in the "holes" in the image. A hole was defined as an area of dark pixels surrounded by light pixels. In this case, it referred to those removed dark areas within apples, such as defects, stem-ends, and calyxes, that needed to be preserved

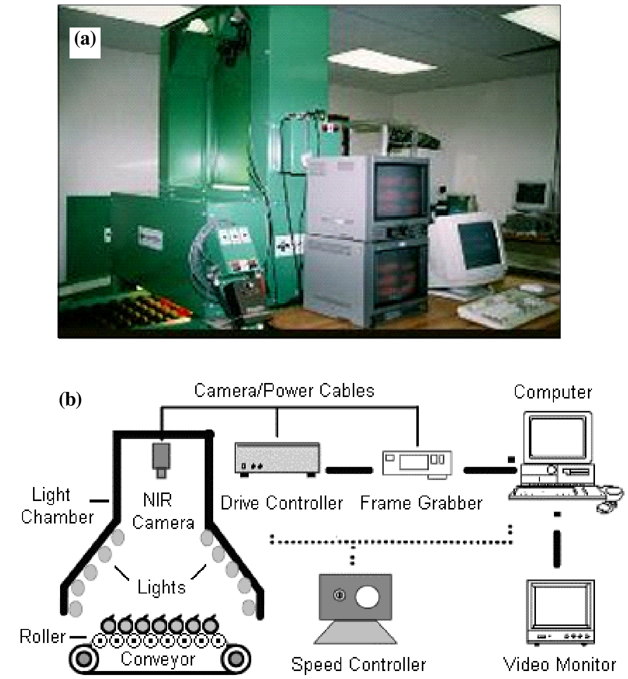


Figure 1. NIR machine vision system for automatic apple sorting and grading: (a) photograph of the system and (b) schematic representation.

for further analysis. Second, an area-open operation was performed, based on pixel counts, that removed the small foreground "objects" with T_2 pixels or less. T_2 was a pre-determined threshold, which was set to 400 in this research. In other words, this operation discarded those brighter pixels that did not belong to the apples but to the background. The discarded pixels came from the false segmentation due to the intensity variation of the background. A small set of ten images was used for threshold T_1 and T_2 training. Then the selected thresholds were applied through testing data set. Because the lighting condition in the imaging chamber was well controlled, image contrast was consistently high from image to image. Therefore, a small training data set was sufficient to determine thresholds T_1 and T_2.

Each apple in a single image also needed to be extracted individually. To achieve this goal, the boundary of each apple was obtained by the aforementioned two-step segmentation. Then coordinates of the circumscribed rectangle of each apple could be easily determined using the boundary information. The segmented images as well as individual apple images can be seen in figures 2b and 2c, respectively.

Estimation of Illuminant Direction

In this study, the direction of the light source was estimated according to the Lambertian reflectance model and the shading information, i.e., 2D image intensity (Pentland, 1982; Zhu et al., 2007b). Given any particular 8-neighborhood direction in the image plane, the following relationship can be obtained:

$$\begin{pmatrix} d\bar{I}_1 \\ d\bar{I}_2 \\ \vdots \\ d\bar{I}_L \end{pmatrix} = \begin{pmatrix} dx_1 & dy_1 \\ dx_2 & dy_2 \\ \vdots & \vdots \\ dx_L & dy_L \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \tag{1}$$

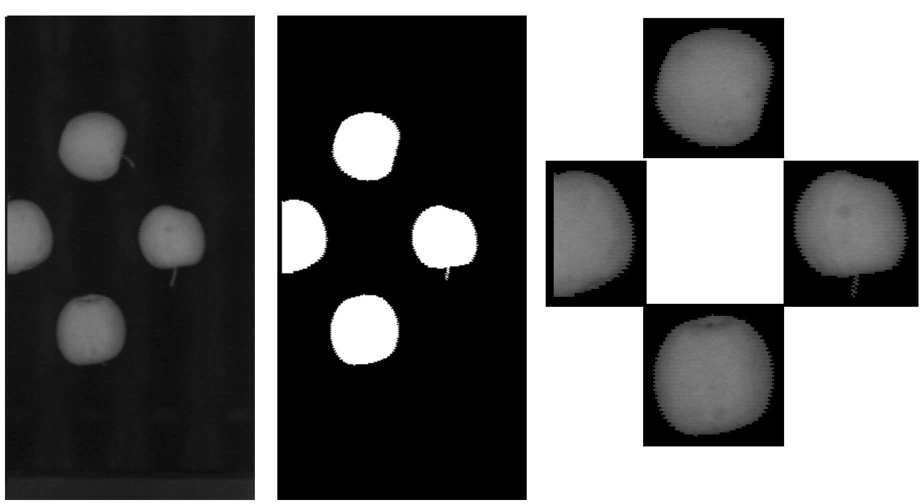
where $d\bar{I}_i$ is the average image intensity change along the ith direction (dx_i, dy_i), (X, Y) is the estimation of the x and y components of the tilt angle τ (the angle that the image plane component of the light source vector makes with the x-axis), and L is the total number of directions considered. In this research, a total of eight directions (0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°, and 360°) were considered. More or fewer angles could be used; eight angles were chosen by balancing the effectiveness and the computation time. When D is defined as the direction matrix in equation 1, the following equation can be obtained:

$$\begin{pmatrix} X \\ Y \end{pmatrix} = (D^T D)^{-1} D^T \begin{pmatrix} d\bar{I}_1 \\ d\bar{I}_2 \\ \vdots \\ d\bar{I}_L \end{pmatrix}, \quad D = \begin{pmatrix} dx_1 & dy_1 \\ dx_2 & dy_2 \\ \vdots & \vdots \\ dx_L & dy_L \end{pmatrix} \tag{2}$$

Then, the tilt angle τ can be determined by:

$$\tau = \arctan\left(\frac{Y}{X}\right) \tag{3}$$

and the slant angle σ (the angle between the illuminant vector and the z-axis) is calculated as:



(a) (b) (c)

Figure 2. (a) Original NIR image of apples, (b) NIR image after two-step segmentation, and (c) extracted individual apple images.

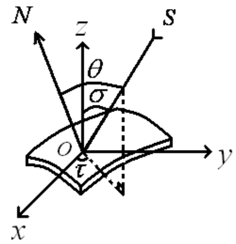


Figure 3. Illustration of Lambertian model.

$$\sigma = \begin{cases} \arccos\left(\sqrt{1-(X^2+Y^2)/K^2}\right) & \text{if } X^2+Y^2 \le K^2 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where $K = \sqrt{E\{dI^2\} - (E\{dI\})^2}$ and the $E\{\bullet\}$ operator is the expectation statistics. The relationship among tilt angle τ, slant angle σ, surface normal *N*, and light source direction *S* is schematically illustrated in figure 3.

Notice that the light settings of our automatic apple imaging system was always fixed and was not changed during the online inspection. Therefore, the illuminant direction was calculated only once and could be done offline before running the system. In this study, ten uniformly arranged warm-white fluorescent lamps were used to provide a uniform lighting condition, which meant that even the hardware configuration was known. The calculation of the light source was still necessary in order to determine the equivalent pseudo light source direction, and hence provide the tilt and slant angle for the shape-from-shading (SFS) approach. Given the Lambertian model as well as the light source information, the SFS method could be used to retrieve the 3D shape of apples based on 2D data.

Shape-from-Shading Model for 3D Apple Surface Reconstruction

Shape-from-shading (SFS) techniques were introduced in the early 1970s (Horn, 1970). They are still widely studied by researchers (Prados et al., 2002; Kimmel and Sethian, 2001; Prados and Faugeras, 2003; Crouzil et al., 2003; Tankus et al., 2004). The basic idea of this approach is to derive a 3D scene description from 2D information, such as a 2D image intensity map. According to Zhang et al. (1999), SFS algorithms can be categorized into four approaches: minimization, propagation, local, and linear. Generally, minimization approaches are more robust, while the other approaches are faster. In this research, Pentland's SFS method (Pentland, 1989), which belongs to the linear category, was selected for the 3D apple surface recovery. Pentland's method was chosen for the following reasons:

- It gave relatively low reconstruction errors under short computation time, which is critical for on-line industrial applications.
- It was discovered that in most of our experiments, given the NIR images of apples, the human eye could identify stem-ends and calyxes according to their 3D information. (The stem-ends and calyxes often had continuously changing intensity due to their concave shape, while the apple defects and other surfaces did not).

The above facts implied that an approach that could simulate the human eye's ability to locate 3D shapes might be feasible for apple stem-end/calyx identification. Because Pentland's method was derived from human eye perception properties and was close to the way human eyes recover 3D information from a 2D scene, it was chosen as the most feasible approach.

In order to introduce the SFS algorithm equation, it is useful to rewrite the light source and surface normal vectors in the following equation:

$$S = (s_x, s_y, s_z) = (\cos\tau\sin\sigma, \sin\tau\sin\sigma, \cos\sigma) \quad (5)$$

and

$$N = (n_x, n_y, n_z) = \frac{(P, Q, 1)}{\sqrt{P^2+Q^2+1}} \quad (6)$$

$$P = \frac{\partial z(x,y)}{\partial x}, \qquad Q = \frac{\partial z(x,y)}{\partial y} \quad (7)$$

Based on the Lambertian model, and Taylor series expansion up to the first order, the following equation can be obtained:

$$I(x,y) = \cos\sigma + P\cos\tau\sin\sigma + Q\sin\tau\sin\sigma \quad (8)$$

By taking the Fourier transform of both sides of equation 8, taking off the DC component, and rearranging the equation, it is easy to get:

$$F_z(f_1, f_2) = \frac{1}{2\pi i}\frac{F_I}{f_1\cos\tau\sin\sigma + f_2\sin\tau\sin\sigma} \quad (9)$$

where $F_z(f_1, f_2)$ is the 2D Fourier transform of the apple depth map $Z(x, y)$, F_I is the 2D Fourier transform of the apple image $I(x, y)$, and f_1 and f_2 are corresponding coordinates in the Fourier domain. Hence:

$$Z(x,y) = IFT\{F_z(f_1, f_2)\} \quad (10)$$

Since both 2D fast Fourier and inverse Fourier transforms are available, the solution can be calculated very quickly.

Quadratic Facet Model for Stem-End and Calyx Convex 3D Shape Fitting

The idea behind the facet model is to view the spatial domain of an image as the combination of connected surface pieces, so-called facets, each of which satisfies certain shape constraints. A sloped/degree-one facet model was employed to fit the test images through least-square estimation (Haralick and Watson, 1981). The fitting results were acceptable. However, some details were lost in the fitted images due to only using the first-order polynomial function in this study. Other 3D shape fitting approaches, including high-order facet models, have also been applied by many researchers in the area of image processing, such as edge detection (Haralick, 1983, 1984; Ji and Haralick, 2002), image segmentation (Besl and Jain, 1988; Lukács et al., 1998), object recognition (Hebert et al., 1995; Blane et al., 2000), and image registration (Scott et al., 1995; Jiang et al., 1992; Wyngaerd and Gool, 2002). In this research, the idea of a facet model was extended to fit the 3D depth data instead of traditional image intensities. By doing so, much better results could be obtained in terms of apple stem-end/calyx identification. A quantitative comparison between 3D depth fitting and image intensity fitting is given in the Results and Discussion section.

Because of the apple's convex shape (concave at the stem-end/calyx) and its smooth surface, it is reasonable to assume that the 3D depth value of a small neighborhood on the apple surface can be approximated by a bivariate quadratic function g, and the canonical form of g can be given by:

$$Z(x,y) \approx g(x,y)$$
$$= k_1 + k_2 x + k_3 y + k_4 x^2 + k_5 xy + k_6 y^2 \quad (11)$$

The above equation can be rewritten based on a set of discrete orthogonal polynomial basis:

$$g(x,y) = \sum_{i=1}^{6} K_i h_i(x,y) \quad (12)$$

where $h_i(x, y) = \{1, x, y, x^2 - 2, xy, y^2 - 2\}$ is a set of orthogonal polynomials, and $x = \{-W, ..., -1, 0, 1, ..., W\}$, $y = \{-W, ..., -1, 0, 1, ..., W\}$ within the small neighborhood, where $2W + 1$ refers to the window size of the neighborhood, and W was set to 2 in this study. By comparing equation 11 with equation 12, it is obvious that:

$$\begin{cases} k_1 = K_1 - 2K_4 - 2K_6 \\ k_i = K_i, \ i = 2,3,..,6 \end{cases} \quad (13)$$

The fitting coefficients K_i can be obtained by projecting the apple 3D surface map onto the orthogonal polynomial basis:

$$K_i = \frac{\sum_{x,y} h_i(x,y) Z(x,y)}{\sum_{x,y} h_i^2(x,y)} = Z \otimes w_i \quad (14)$$

where

$$w_i = \frac{\sum_{x,y} h_i(x,y)}{\sum_{x,y} h_i^2(x,y)} \quad (15)$$

where $\otimes$ is the convolution operator. The fitting coefficient K_i is computed by convolving the 3D surface map with the corresponding weight kernel w_i (Ji and Haralick, 2002), which makes the computations much easier than calculating the k_i values from equation 11 directly.

Among computed fitting coefficients, only K_4 and K_6 are needed to describe the shape of a quadric surface of a given type, while the other coefficients are used to control the orientation and translation of the surface (Besl and Jain, 1985). Some typical quadric shapes affected by coefficients K_4 and K_6 are illustrated in figure 4.

In this research, the orientation of the facet was negligible compared to its shape; meanwhile, the translation of the facet has already been taken care of during the convolution. Therefore, identification of the concave shape on the convex apple surface could be achieved by simply checking coefficients K_4 and K_6 under a pre-determined threshold. K_4 and K_6 were chosen according to how they affect the concavity of a 3D surface. Both parameters were set to zero (a very general threshold) and kept fixed during entire experiment for 203 sample images. This showed that the detection rate of stem-end/calyx identification was not very sensitive to changes in these coefficients. This was expected, since the difference between 3D concave and convex shapes is quite significant. Therefore, only qualitative analysis of the surface shape was necessary for apple stem-end/calyx identification. In other words, it was only necessary to know whether the fitted shape was concave or not, and not exactly how concave the shape was. In addition, this was a simplified version of 3D surface fitting, and it was easier to perform than other 3D fitting approaches. (Simplicity is always preferred by the industry, since it will save time and hence increase throughput.) Similarly, the threshold was not very vulnerable to changes in lighting settings as well as apple varieties, since the difference between an apple 3D surface and its concave stem-end/calyx would always remain the same. However, if a lighting change caused inaccuracy of apple 3D reconstruction, then the detection accuracy could be affected.

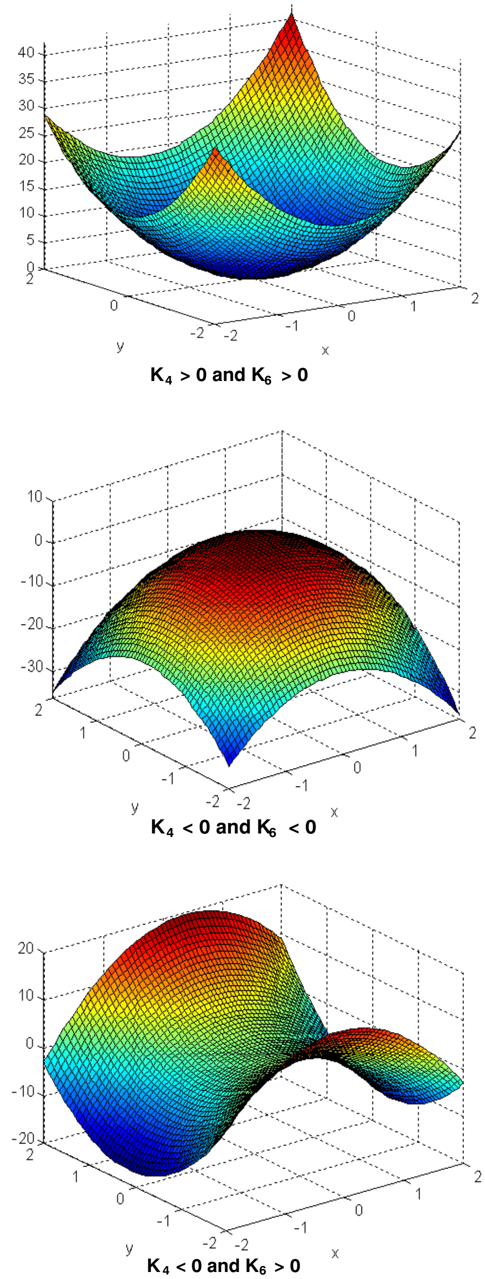


Figure 4. Surface shape affected by K_4 and K_6.

The procedure for using the quadratic facet model for apple stem-end/calyx convex 3D shape fitting is summarized as follows:

1. Estimate the azimuth angle τ and the slant angle γ according to equations 3 and 4.
2. Compute the 3D depth map of the original NIR apple image I according to equations 9 and 10.
3. Decompose the 3D depth map of the NIR apple image into quadratic format according to equations 11 through 14 and obtain the quadratic coefficients Ki.
4. Threshold Ki to distinguish concave shapes on the convex apple surface in order to extract the apple stem end and calyx.

Because the proposed approach differentiated apple stem-ends/calyxes from defects based on their different 3D shapes, it gave a better performance than the classification methods that only used 2D information such as image intensity and shape.

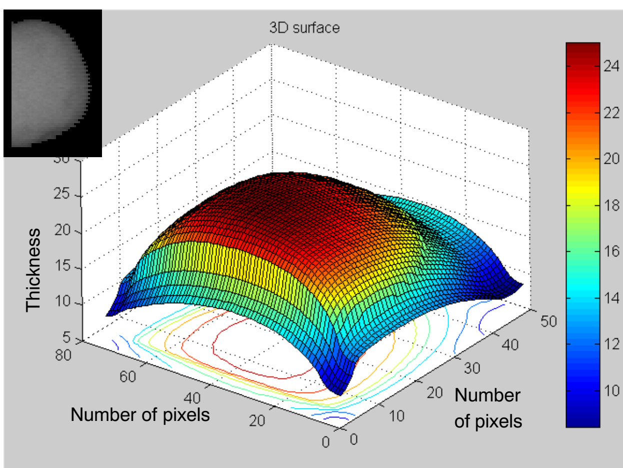
Results and Discussion

3D Surface Reconstruction

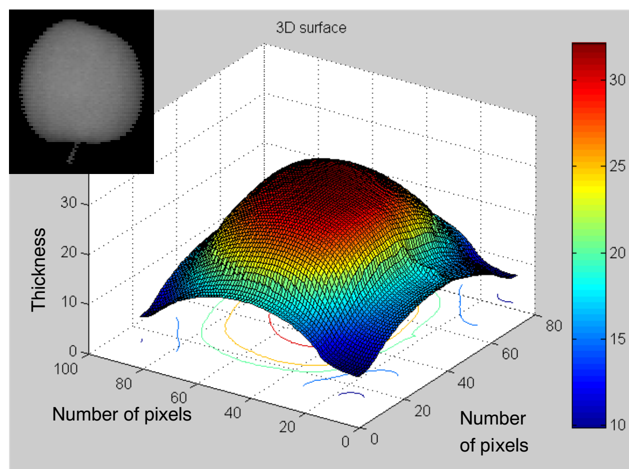
Reconstructed 3D surface maps of Golden Delicious apples using equation 10 are shown in figure 5, which includes five groups (two images per group) based on different apple/image conditions. As shown in figure 5, 3D apple surfaces were successfully recovered from 2D NIR images (shown at the top-left corner of each image). Given insufficient data, such as half or part of the apple image, the 3D map of half apples in (a) and (g) could still be restored without any visible distortion. Different 3D shape properties among normal apple surfaces, stem-ends/calyxes, and defects can be observed from (c) to (h). Generally, the normal apple surface was a convex 3D shape, while the defects exhibited small indentations, which were much shallower and flatter in their 3D depth than the concave shape of the stem-ends and calyxes. The deep concave shape of the stem-ends/calyxes made it possible for the facet model to detect their correct positions on the apple surface. Corrupted image data were also tested to show the robustness of the algorithm. Relatively good results were achieved, as shown in (i) and (j). Once the depth value for apple stem-ends/calyxes was obtained, localization of the stem-end/calyx could be achieved according to the quadratic facet model. The identification results are given in the next section.

3D Quadratic Facet Fitting

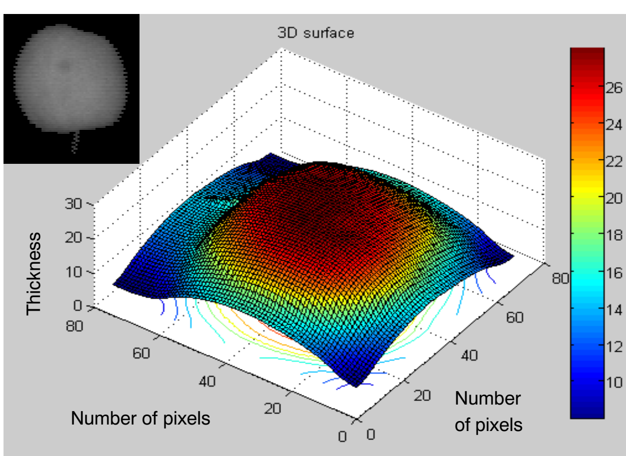
Figure 6 shows how the stem ends and calyxes can be detected by the facet fitting approach using 3D depth data. Fig–



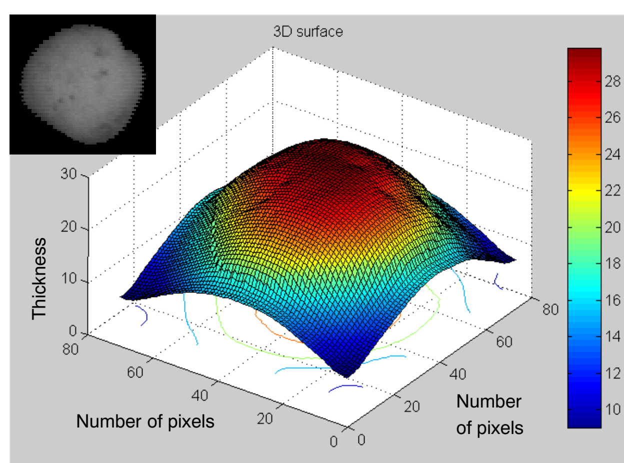
(a) Normal half apple



(b) Normal whole apple

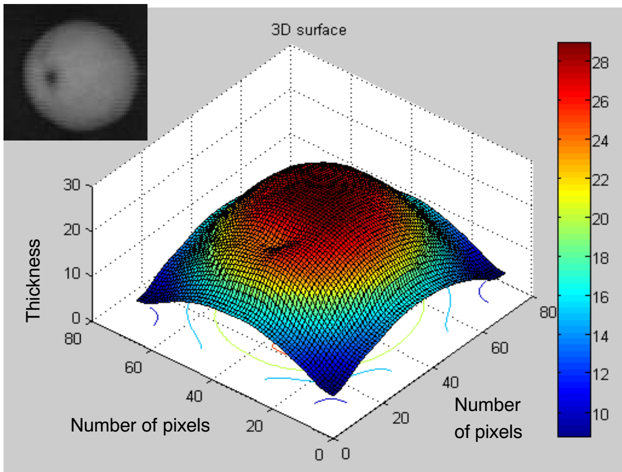


(c) Apple with single defect

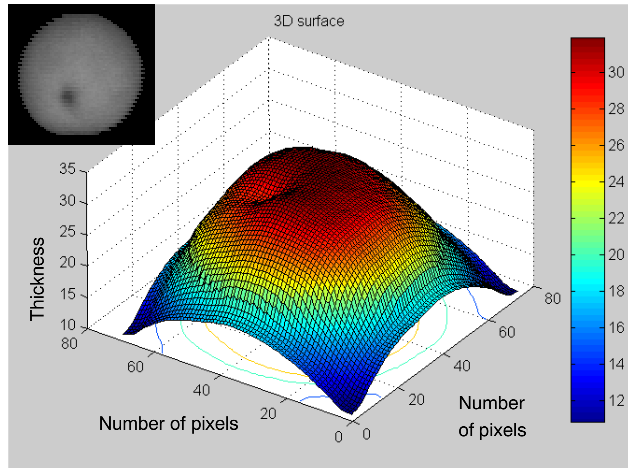


(d) Apple with multiple defects

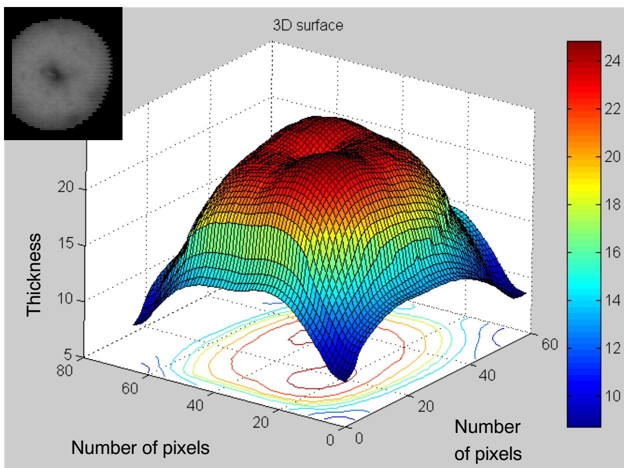
Figure 5. **(continued on next page)**



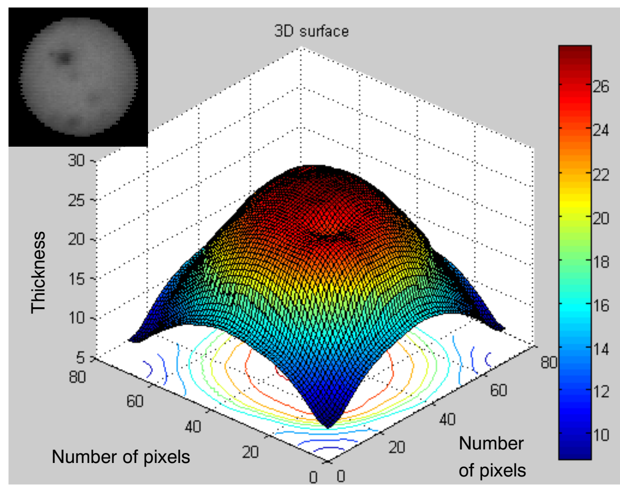
(e) Apple with calyx



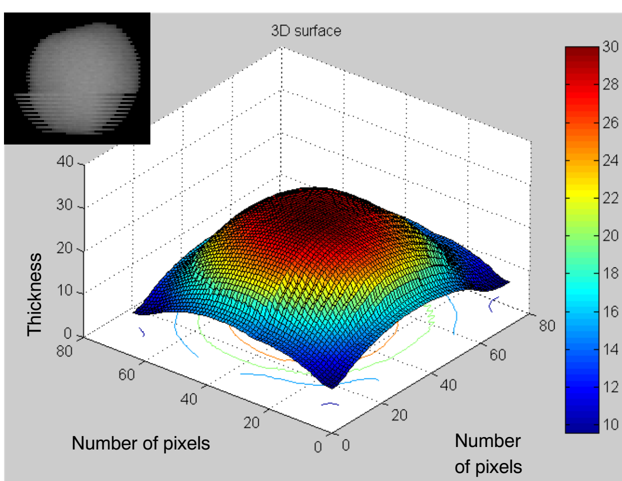
(f) Apple with calyx



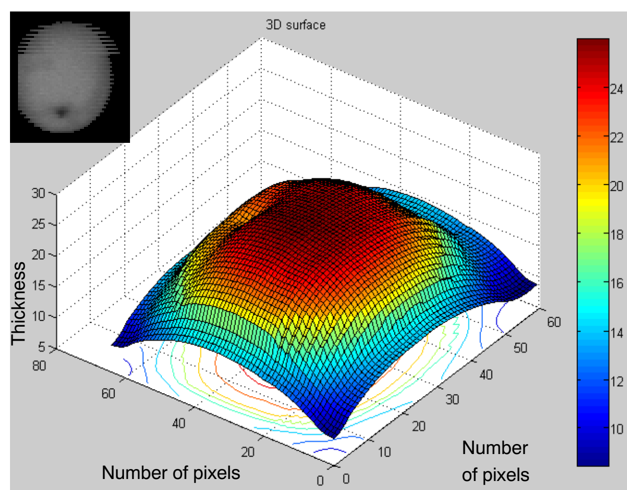
(g) Apple with stem/calyx and defects



(h) Apple with stem/calyx and defects



(i) Corrupted image data (normal apple)



(j) Corrupted image data (apple with calyx)

Figure 5. Reconstructed 3D surface maps of five groups of Golden Delicious apples (two images per group) based on different apple/image conditions. Note: The *z*-axis indicates relative depth. During the reconstruction of the apple 3D surface, DC components and high-order Taylor expansion were removed. Therefore, the unit of the *z*-axis is actually undefined. However, this will not affect the detection results since only relative depth of the apple 3D surface was considered.

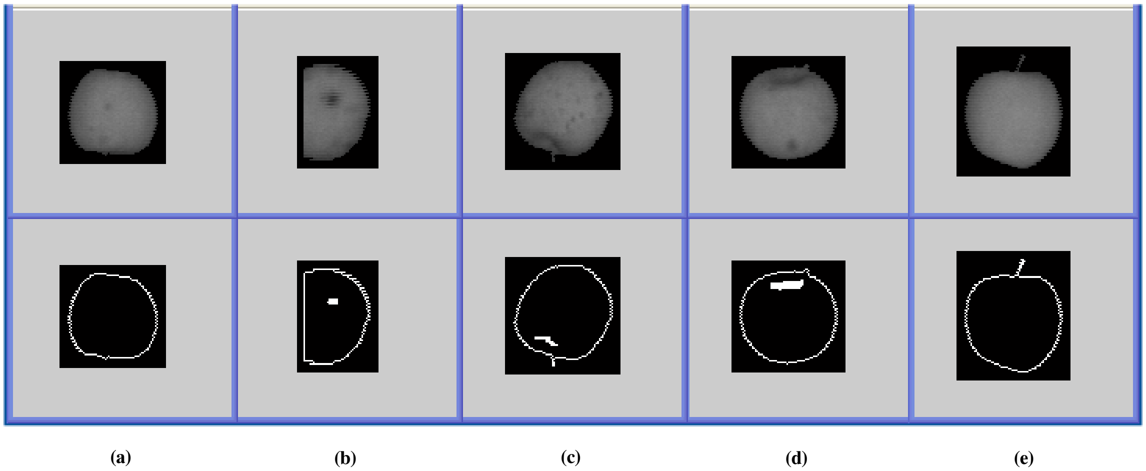


Figure 6. Quadratic facet fitting using apple 3D depth map.

ure 6a shows that several small defects on the apple, but no stem-ends/calyxes, are facing the camera. Figure 6b shows a 2/3 apple image with a calyx at the center of the apple. The calyx was located correctly by the 3D fitting model. Figure 6c shows an apple with many defects, and its stem-end is facing down. The fitting approach successfully detected the location of that stem-end, while leaving all defects unlabeled. This can also be seen in figure 6d, although the defect on the bottom of the apple is large and dark. The fitting model correctly distinguished the stem-end from that defect according to their different 3D properties. Figure 6e shows a good apple sample.

Unlike the traditional facet fitting approaches, which usually use image intensity value as the fitting input, this study used the recovered 3D depth as the fitting data. In other words, the image intensity $I(x, y)$ was substituted by the 3D depth $Z(x, y)$ in equation 14. By doing so, much better results could be achieved. A comparison between these two methods is shown in figure 7, which shows the original and processed NIR images. The first row shows the original images, the second row shows the stem-end/calyx identification results using intensity data as the input, and the third row represents the identification results using 3D depth as the input. Figure 7a is a good apple sample, but one false alarm was generated by the traditional fitting method. One calyx in figure 7b was misidentified by the traditional method, and a false alarm can be found at the same time. In addition, many apple defects were misclassified as stem-end/calyx by the traditional method, which can be seen in figures 7c and 7d. For all five examples in figure 7, the method presented in this study gave the correct identification results.

A total of 203 NIR Golden Delicious apple images were tested. The detailed composition and statistics of the data are shown in table 1. Given three different test criteria, a consistent detection rate was achieved by the 3D depth fitting method. All detection rates were greater than or equal to 90%.

A comparison between the traditional fitting method and 3D fitting method is also demonstrated in figure 8. Both type I and type II errors (Ott and Longnecker, 2001) were considered in the study to evaluate the performance of the proposed

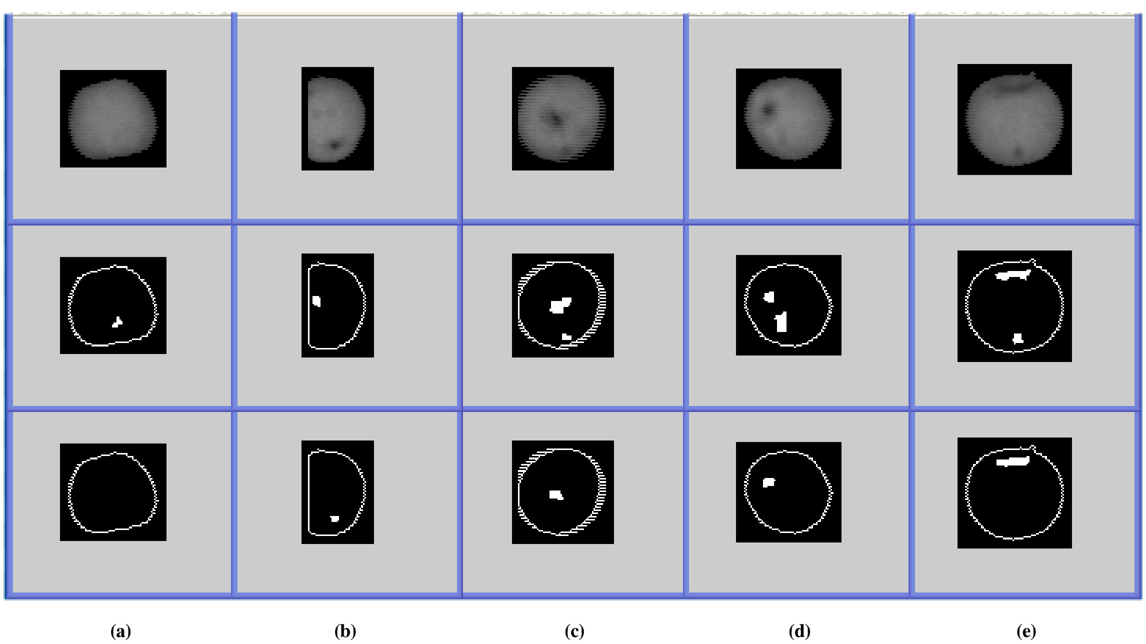


Figure 7. Comparison on fitting results using original image intensity (second row) and 3D depth (third row).

Table 1. Data composition and detection rate of 203 test samples.

Category	No. of Samples	No. of Samples Detected	Detection Rate (%)
Samples with stem end or calyx facing the camera	140	126	90.00
Samples without stem end or calyx facing the camera	63	57	90.48
Total samples	203	183	90.15

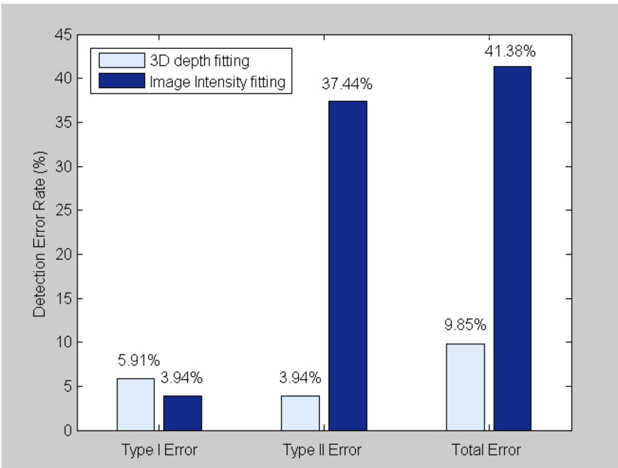


Figure 8. Comparison between 3D depth and image intensity fitting.

approach. Type I error was calculated as the number of incorrectly classified samples (i.e., defected apple images) divided by the total number of samples, while type II error was computed as the number of false classified samples over the total number of samples. Although the type I error of the first method was slightly lower than that of the second approach, a much lower type II error was obtained by the second method. As a result, the overall error rate was reduced significantly (from 41.38% to 9.85%) by using the 3D depth fitting method. In addition, the computation time of the proposed algorithm was tested according to all 203 apple images, and an average of 0.042 ms per image was achieved using a moderate PC configuration (Dell Dimension 4600C), which included a Pentium 4 2.4 GHz CPU and 256 MB of RAM. Matlab Version 7.0.0 was used as the programming software.

Conclusions

In this study, apple 3D surface reconstruction was achieved based on an SFS algorithm. Unlike structured light range imaging, which uses only partial information of the apple surface, this approach took advantage of the complete image information. Every pixel value contributed to the reconstructed 3D map, which meant that a more detailed 3D description could be obtained. In addition, because the camera used in this study operated in interlaced scanning mode, a zigzag effect around the apple boundary was inevitably generated due to the high conveyer speed (see fig. 1). However, this method showed its robustness and still worked well regardless of this effect. The results can be improved if a high-resolution, high-speed progressive scan camera is used in the future. There was no additional light source required in the system; normal visible white light plus an NIR filter was sufficient. There was also no need to image a whole apple; just half of an apple in the camera's field of view could be recovered without any distortion. Given successfully recovered 3D depth data, a quadratic facet model was followed to locate the apple stem-ends/calyxes based on their 3D properties. A total of 203 Golden Delicious apple images were tested, and an average 90.15% detection rate was achieved. A comparison between a traditional facet fitting method and the 3D depth fitting method showed that the latter approach performed much better than the traditional one. The method used in this study obtained a 9.85% total error rate compared to 41.38% when using the traditional method. This also meant that the 3D surface model was effective for apple stem-end/calyx identification.

In this research, both the experimental results and the comparison with traditional 2D surface fitting method showed the effectiveness of the proposed 3D based approach. However, a quantitative analysis of the accuracy for SFS 3D recovery is still necessary in future study. Furthermore, more test samples will be needed to evaluate the robustness of the proposed method. Some other surface reflection models as well as 3D fitting models might also be considered to further exploit the application of apple stem-end/calyx identification.

Acknowledgements

The authors are grateful to Dr. Qinfen Zheng at the University of Maryland for his valuable suggestions. Thanks are also due to Ms. Abby Vogel for her critical reading of the manuscript.

References

Bennedsen, B. S., D. L. Peterson, and A. Tabb. 2005. Identifying defects in images of rotating apples. *Comput. Electron. Agric.* 48(2): 92-102.

Besl, P. J., and R. C. Jain. 1985. Three-dimensional object recognition. *ACM Computing Surveys* (*CSUR*) 17(1)**:** 75-145.

Besl, P. J., and R. C. Jain. 1988. Segmentation through variable-order surface fitting. *IEEE Trans. PAMI* 10(2): 167-192.

Blane, M. M., Z. Lei, H. Civi, and D. B. Cooper. 2000. The 3L algorithm for fitting implicit polynomial curves and surfaces to data. *IEEE Trans. PAMI* 22(3): 298-313.

Brown, G. K., L. J. Segerlind, and R. Summit. 1974. Near-infrared reflectance of bruised apples. *Trans. ASAE* 17(1): 17-19.

Cheng, X., Y. Tao, Y. R. Chen, and Y. Luo. 2003. NIR/MIR dual-sensor machine vision system for on-line apple stem-end/calyx recognition. *Trans. ASAE* 46(2): 551-558.

Crouzil, A., X. Descombes, and J.-D. Durou. 2003. A multiresolution approach for shape from shading coupling deterministic and stochastic optimization. *IEEE Trans. PAMI* 25(11): 1416-1421.

Good Fruit Growers. 1993. MERLIN electronic color sorter. *Good Fruit Growers* (March 15).

Haralick, R. M. 1983. Ridges and valleys on digital images. *Comput. Vision Graph. Image Proc.* 22: 28-38.

Haralick, R. M. 1984. Digital step edges from zero crossing of second directional derivatives *IEEE Trans. Pattern Analysis and Machine Intelligence* 6(1): 58-68.

Haralick, R. M., and L. Watson. 1981. A facet model for image data. *Comput. Graphics Image Proc.* 15: 113-129.

Hebert, M., K. Ikeuchi, and H. Delingette. 1995. A spherical representation for recognition of free-form surfaces. *IEEE Trans. PAMI* 17(7): 681-690.

Horn, B. K. P. 1970. Shape from shading: A method for obtaining the shape of a smooth opaque object from one view. PhD diss. Cambridge, Mass.: Massachusetts Institute of Technology.

Ji, Q., and R. M. Haralick. 2002. Efficient facet edge detection and quantitative performance evaluation. *Pattern Recog.* 35(3): 689-700.

Jiang, H., R. A. Robb, H. Tainter, and S. Kerrie. 1992. New approach to 3-D registration of multimodality medical images by surface matching. *Proc. SPIE* 1808: 196-213.

Jing, H., and Y. Tao. 1999. Real-time high-resolution 3-D laser range imaging of poultry products. ASAE Paper No. 993142. St. Joseph, Mich.: ASAE.

Kimmel, R., and J. A. Sethian. 2001. Optimal algorithm for Shape from Shading and path planning. *J. Math. Imaging and Vision* 14(3): 237-244.

Leemans, V., H. Magein, and M.-F. Destain. 1999. Defect segmentation on 'Jonagold' apples using colour vision and a Bayesian classification method. *Comput. Electron. Agric.* 23(1): 43-53.

Li, Q., M. Wang, and W. Gu. 2002. Computer vision based system for apple surface defect detection. *Comput. Electron. Agric.* 36(2-3): 215-223.

Lukács, G., R. Martin, and D. Marshall. 1998. Faithful least-squares fitting of spheres, cylinders, cones, and tori for reliable segmentation. *Lecture Notes in Comput. Sci.* 1406: 671.

Ott, L., and M. Longnecker. 2001. *An Introduction to Statistical Methods and Data Analysis.* 5th ed. Pacific Grove, Cal.: Wadsworth Group.

Penman, D. W. 2001. Determination of stem and calyx location on apples using automatic visual inspection. *Comput. Electron. Agric.* 33(1): 7-18.

Pentland, A. P. 1982. Finding the illuminant direction. *J. Opt. Soc. America A* 72(4): 448-455.

Pentland, A. 1989. Shape information from shading: A theory about human perception. *Spatial Vision* 4(2/3): 165-182.

Prados, E., and O. Faugeras. 2003. "Perspective shape from shading" and viscosity solutions. *Proc. 9th IEEE Intl. Conf. on Comput. Vision* 2: 826-831. Piscataway, N.J.: IEEE.

Prados, E., O. Faugeras, and E. Rouy. 2002. Shape from shading and viscosity solutions. *Proc. 7th European Conf. on Comput. Vision* 2: 790-804. Berlin, Germany: Springer.

Scott, A. M., H. Macapinlac, J. Zhang, F. Daghighian, N. Montemayor, H. Kalaigian, G. Sgouros, M. C. Graham, K. Kolbert, S. D. J. Yeh, E. Lai, S. J. Goldsmith, and S. M. Larson. 1995. Image registration of SPECT and CT images using an external fiduciary band and three-dimensional surface fitting in metastatic thyroid cancer. *J. Nuclear Med.* 36(1): 100-103.

Shahin, M. A., E. W. Tollner, R. W. McClendon, and H. R. Arabnia. 2002. Apple classification based on surface bruises using image processing and neural networks. *Trans. ASAE* 45(5): 1619-1627.

Tankus, A., N. Sochen, and Y. Yeshurun. 2004. Perspective shape-from-shading by fast marching. *Proc. IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR '04*) 1: 43-49. Piscataway, N.J.: IEEE.

Tao, Y. 1996. Spherical transform of fruit images for on-line defect extraction of mass objects. *Opt. Eng.* 35(2): 344-350.

Tao, Y., and Z. Wen. 1999. An adaptive spherical image transform for high-speed fruit defect detection. *Trans. ASAE* 42(1): 241-246.

Throop, J. A., D. J. Aneshansley, B. L. Upchurch, and B. Anger 2001. Apple orientation on two conveyors: Performance and predictability based on fruit shape characteristics. *Trans. ASAE* 44(1): 99-109.

Unay, D., and B. Gosselin. 2004. An approach for recognizing stem-end/calyx regions in apple quality sorting. In *Proc. Advanced Concepts for Intelligent Vision Systems Conf.* (*ACIVS 2004*). Brussels, Belgium.

Wen, Z., and Y. Tao. 1998a. Fuzzy-based determination of model and parameters of dual-wavelength vision system for on-line apple sorting. *Opt. Eng.* 37(1): 293-299.

Wen, Z., and Y. Tao. 1998b. Brightness-invariant image segmentation for on-line fruit defect detection. *Opt. Eng.* 37(11): 2948-2952.

Wen, Z., and Y. Tao. 1999. Building a rule-based machine-vision system for defect inspection on apple sorting and packing lines. *Expert Sys. Appl.* 16(3): 307-313.

Wen, Z., and Y. Tao. 2000. Dual-camera NIR/MIR imaging for stem-end/calyx identification in apple defect sorting. *Trans. ASAE* 43(2): 449-452.

Wyngaerd, J. V., and L. V. Gool. 2002. Coarse registration of surface patches with local symmetries. *Lecture Notes in Comput. Sci.* 2351: 572.

Yang, Q. 1996. Apple stem and calyx identification with machine vision. *J. Agric. Eng Res.* 63(3): 229-236.

Yang, Q., and J. A. Marchant. 1996. Accurate blemish detection with active contour models. *Comput. Electron. Agric.* 14(1): 77-89.

Zhang, R., P.-S. Tsai, J. E. Cryer, and M. Shah 1999. Shape from shading: A survey. *IEEE Trans. Pattern Analysis and Machine Intelligence* 21(8): 690-706.

Zhu, B., L. Jiang, X. Cheng, and Y. Tao. 2005. 3D surface reconstruction of apples from 2D NIR images. *Proc. SPIE* 6000: 242-251.

Zhu, B., L. Jiang, Y. Luo, and Y. Tao. 2007a. Gabor feature-based apple quality inspection using kernel principal component analysis. *J. Food Eng.* 81(4): 741-749.

Zhu, B., L. Jiang, and Y. Tao. 2007b. 3D shape enhanced transform for automatic apple stem-end/calyx identification. *Opt. Eng.* 46(1): 17201.1-9.

Zhu, B., L. Jiang, and Y. Tao. 2007c. 3D Surface reconstruction and analysis of apple near-infrared data for the application of apple stem-end/calyx identification. ASABE Paper No. 073074. St. Joseph, Mich.: ASABE.

Zion, B., P. Chen, and M. J. McCarthy. 1995. Detection of bruises in magnetic resonance images of apples. *Comput. Electron. Agric.* 13(4): 289-299.